Reproduced from
Studio Sound, May 1974

What's wrong with quadraphonics

Michael Gerzon

The BBC recently held a press conference on four channel sound to tell us that they had nothing to report. With the rest of the industry wondering what happens next, and in some cases not caring, Michael Gerzon suggests remedies for defects in presently available systems.

Few people in the recording business (and, if it comes to that, few record buyers) are happy with the present state of the quadraphonic art. Those who are have argued that this is a temporary state of affairs (like the hostility which met the first days of stereo) which will evaporate as soon as the quadraphonic skills of the recording engineer increase with experience. As one who has argued for surround sound from the beginning, and whose first experiments were in 1965 and 1968, may I beg to differ? Quadraphonics, as at present widely conceived, is a dead end.

Don't misunderstand this flat pronouncement. I am not saying there is no future for surround sound in the home via four or more speakers. What I am saying is that most present methods of achieving this aim, whether they be 'matrix' or 'discrete', are quite incapable of optimum results and do not do what they claim to do. The fault lies partly in studio equipment incapable of giving good quadraphonic results, and partly in erroneously conceived 'quadraphonic' systems which leave the apparent localisation of sounds at the mercy of the listener's imagination.

The aim of quadraphonic systems has in the main been to duplicate the effect of the 'original' four track mastertape in the recording studio. This would be fine if one were sure that the four track tape were actually duplicating the precise intentions of the recording engineer and producer. The trouble is: it isn't.

Any producer is at the mercy of the limitations of the studio equipment. If he wishes for a certain effect on the recording, he may well have to accept a poor imitation of it if the studio equipment is not capable of the precise effect he wants. Moreover, the effect desired may only be obtained even approximately at precisely one listening position with one size and shape of loudspeaker layout, with the listener having to face precisely in one direction. Clearly a recorded effect that is so critical is not much use except for indulging in ego-trips.

And yet some of the effects obtained by conventional 'discrete' systems are precisely of this nature. I refer in particular to side-centre sounds (*ie* side-right or side-left). These are conventionally obtained by feeding or panning a mono sound equally to the two speakers on the relevant side. It is a matter of experience that a forward-facing listener does not hear such a sound as lying completely at his side unless he seats himself in *just* the right position and takes care not to move his head even by the tiniest amount. Otherwise the sound tends to jump to the front or to the back speaker at that side with the slightest provocation. Some producers of quadraphonic drama have refused to use side positions because of this unpredictability, and one inventor of 'matrix' systems has justified the admittedly dreadful side performance of his 'matrix' system by pointing out that sound engineers 'do not want to use side positions anyway'. Of course they don't, if the results are bad!

Sounds positioned in the front and back quadrants with conventional quadraphonic panpotting are only a little better. In the bad old days of two-speaker stereo, we were all told how bad it was to angle speakers more than 60^{o} apart at the listening position, as one then got 'hole-in-the-middle', where middle sounds fled to one edge or the other of the image at the slightest provocation. Yet now some people expect good results from 'discrete' quadraphony with the front speakers angled 90^{o} apart. The laws of good sound haven't changed and, if one wants stable sound images between the speakers, one is forced to sit in a 'quadraphonic seat' that is tiny in comparison even to the 'stereo seat'.

The effects of this grotesquely poor sound image stability have been predictable. Since the only 'discrete' sound images that are stable are the four corners, sounds have been piled up in those positions with gaping holes left elsewhere. Since even front centre images are unstable, there has been a temptation to indulge in (legitimate) 'interior' positioning effects which sound like no position in the real world but which have the virtue of not being any more 'wrong' in one listening position than in another. In so far as between-speaker sound images have been made to work, this has been achieved mainly by the ears noticing a difference in directional effect from corner-only sounds. When between-speaker sounds occur on their own, they become hard to localise.

Of course, there can be no disputing that recordings based mainly or solely on the four corners are capable of a great artistic success – one notes the acknowledged masterpieces of electronic quadraphonic music *Philomel* by Milton Babbitt and *Kontakte* by Stockhausen. However, one cannot deny that not to have any other possibilities is extremely restrictive, both for the creative Pop or electronic producer and for the classical man with a 'concert hall' approach.

The poor localisation of 'discrete' four channel tapes made using conventional quad panpots is made even worse by two other phenomena that have been known for many years. It was shown by de Boer in 1947[1] that widely-angled stereo speakers produced a sound image which, after some experience, could be interpreted as being elevated above or depressed below the line joining the speakers. The effect was small for speaker angles up to 60^{O} but the elevation increased to around 40^{O} for a 90^{O} interspeaker angle. This effect was not always heard by inexperienced listeners, which suggests that the directional information reaching the ear is not heard as a 'natural' sound position but that one can learn to give it some sort of interpretation. In fact, the effect is closer to the 'interior' effect than to that of height.

Shuffler circuit

The other phenomenon disturbing localisation is the effect described in 1957 by Clarke, Dutton and Vanderlyn[2] in connection with the EMI 'Stereosonic' system, whereby the width of a stereo image in the treble differed from that in the bass. They suggested the use of a 'shuffler' circuit which narrowed the treble by reducing treble stereo separation. However, this does not reduce the degree of image blurring thus caused, as shown by Harwood[3] in 1968. This blurring is not too disturbing when confined to a mere 60^{O} of angle, but is not acceptable in a 360^{O} image.

Thus far there have been three published approaches to surround sound reproduction. Starting with the worst approach and ending with the best, these are:

(1) 'Matrix' systems, which aim to imitate 'discrete' systems via less than four channels.

(2) 'Discrete' systems, which use four channels and create phantom inter-speaker images by feeding (panning) sounds only to the two adjacent speakers.

(3) 'Harmonic Synthesis' or 'kernel' systems.

The last approach requires some explanation. Commercial examples are the UMX family of systems of Nippon Columbia and the Japanese RM Regular Matrix system (which does *not* include Sansui's QS system which is only an approximation to RM), and also the British NRDC ambisonic system.

Kernel systems start from the observation that the effects one would like to produce include a continuum of directions around the listener. Such systems imagine a limited number of channels being used to convey the sound to the listener, but are designed to recreate a continuous range of directions around the listener approximating the original. This recreation may take place via (say) only four speakers but the signals fed to the speakers do not in themselves matter; only the directional effect of the sound field at the listener matters. (It is interesting that this philosophy is close to that expressed in Blumlein's famous 1931 stereo patent.[4])

Kernel algebra

Because the aim of a kernel system is to convey through a finite number of channels an infinite number of directions (and thus an infinite number of channels), the mathematics used is not 'matrix' algebra (which is used only to describe transformations of a finite number of variables) but 'kernel' algebra (which is the corresponding mathematics used when one has an infinite continuum of variables).

Although the author was working with a kernel approach to surround sound as early as 1970[5], and had privately developed a general mathematical theory of such systems at the time, such systems were developed rather later than 'discrete' approaches (1968) or 'matrix' approaches (1969). As a result, such kernel systems are only now beginning to be marketed, by Nippon Columbia and also as the NRDC ambisonic system (with which the author has recently become associated).

Properly designed kernel systems are capable of results considerably superior to 'discrete' systems[6], and even a two channel kernel system gives results comparable to a four-channel discrete system. This is achieved by feeding signals to all four speakers to create phantom images, and not just the two speakers adjacent to the desired sound position. If one still thinks in terms of 'discrete' systems, the signals fed to the other speakers would be called 'crosstalk' but this crosstalk is not undesirable. Quite the opposite, it is absolutely vital in order to ensure the correct localisation of the phantom sound image.

Several difficulties lie in the way of adopting kernel systems. The most important is the present lack of availability of studio equipment for handling it. It is perfectly feasible to design kernel-type panpots but, as far as I am aware, none has yet been marketed. Such panpots would indeed give side-position sounds when they say they do, and would give front or back quadrant sound with stable localisation. A detailed study shows that kernel systems are capable of a far wider range of control facilities for creative or realistic studio use than is any 'discrete' approach. Devices are being developed at the moment that permit sounds to be moved close to one's ear or far away, which rotate whole sound fields, which pan sounds in the whole of three-dimensional space (above and below as well as in all horizontal directions), and which modify the spatial distribution of sounds without sacrificing good sound imaging. Most of these things cannot even be approximated by 'discrete' or 'matrix' approaches.

Compatible system

It is possible to convey kernel recordings in a reasonably compatible manner via many existing 'quadraphonic' media, including RCA/CD4 type 'discrete' discs, four channel tapes, cartridges and cassettes, RM discs, the Nippon/Columbia UMX systems, and via two channel, three channel or Dorren FM. The only system directly incompatible with a 'kernel' or sound-field approach is the SQ system which needs a vari-matrix type interface unit to convert kernel recordings to SQ.

However, there is a yet more advanced approach based on the kernel approach but designed to ensure the best possible results. I call this the 'Psycho-acoustic' approach because it is based on the idea of tailoring the sound field at the listener to give results subjectively as close as possible to the original effect picked up by the microphones or intended by the producer. Conventionally, the way of inventing quadraphonic systems has been to state a few desired mathematical conditions and then to adjust the decoding unit empirically to give the 'best' possible result with the fixed encoding system thus obtained. This applies to most kernel systems as much as to discrete and matrix systems. The trouble is that there are tens or hundreds of variables in the decoder that can be adjusted, and each decoder has to be tested on a wide variety of sound material. Thus it would be most surprising if the best systems are actually found by this method.

As far as I am aware, systems now under development under the direction of the National Research & Development Corporation are the first to have been developed the other way round. What we have done is to study the various mechanisms by which the ears localise sounds. This, after all, is what we want to get right in the listener's home! A considerable number of calculations of the sound localisations deduced by various possible theories were made for various different systems of reproduction. Rather than assuming one particular theory, those systems are isolated that gave the best possible results according to all of a number of theories. These theories[10] were chosen to fit experimental localisation data by a wide variety of workers. Thus, rather than guessing a system and then trying to make it work, we said what we wanted it to do according to available experimental evidence on the human ears and constructed systems which would do this.

The results showed that in fact several of the available kernel systems were probably capable of being reproduced with good results *provided* that special decoders different from those hitherto suggested were used. Among systems that passed the theoretical tests were the regular matrix systems, the two-channel (BMX) and three channel (TMX) versions of the UMX system, but *not* (via four speakers) the four channel QMX version. Other failures were the two channel periphonic (*ie* with-height) system of Peter Schieber[7] and the author[8], and also the four channel periphonic or tetrahedral systems[5,8,9], when played via a tetrahedron of speakers. The anomalies of the latter system were calculated to come rather close to those experienced by the author in earlier experimental work[9]. However, theory shows that the anomalies disappear when a cubic speaker layout is used for periphony, and this has been confirmed experimentally by John Wright.

A surprising discovery has been a new three channel periphonic system which satisfied none of the mathematical criteria laid down by the author in his general theoretical paper[8] on with-height reproductions but passes the 'psycho-acoustic' theoretical tests with flying colours. This system is suitable for FM broadcasting and fits neatly in the universe of workable systems. It has many desirable properties which make it distinctly attractive, and it now seems possible to design periphonic systems that are usable in the home.

The situation now, as far as studio practice is concerned, is that recordings can be made using four channels that may be encoded and decoded via any of the technically feasible domestic surround-sound systems that are capable of good psycho-acoustic results. However such four channel kernel recordings require other studio processing devices (panpot, microphones etc) than those presently commercially available. Moreover, the four channel signals thus obtained do not, of themselves, give ideal surround sound reproduction but must be fed to decoder circuits that are advanced variants of the primitive and ineffective 'shuffler' circuit used in the Stereosonic system[2,3]. Both for reasons of mono and stereo compatibility and for ease of processing, the decoded version of the signal cannot be used in the studio while the signal is being processed or dubbed, but may only be used for monitoring purposes.

A-format

We thus end up with four types of four channel 'surround sound' signal in the history of psycho-acoustic kernel recording. It starts off life in what I term A-format, which is the form in which the four channel signal is derived from the microphones, taped, mixed, pan-potted etc. This signal format has been chosen to be reasonably compatible with existing 'discrete' four channel signal practice, and uses L_B, L_F, R_F, R_B signals as usual. Another signal format, also not involving any phase shifts, is used within studio processing equipment for reasons of simplifying design. This is known as B-format. The signal format on disc, tape or radio by which the information is conveyed (via two, three or four channels) to the consumer is called C-format ('coded' format) and may involve phase shifts to ensure compatibility or other desirable properties. Lastly, we have D-format ('decoded' format) which is the set of signals fed to the listener's loudspeakers to produce a correct subjective effect in the listening area. The D-format signals can be chosen in a fashion largely independent of the C-format coding used to convey the information to the consumer.

At first this seems an absurd complexity but it should not in fact make life difficult. For most of his working life the recording engineer will handle A-format and only the circuit designer need know about B-format which is a sort of four channel 'sum-and-difference' signal. Moreover the conversion circuitry is simple and the formats have been chosen for minimum complexity and maximum compatibility with present-day 'discrete' signals so that no sudden revolution in the recording methods will be required. D-format will only be needed in the studio for monitoring and is necessary if the best and most accurate reproduced effect is required. For a rather less accurate effect, A-format may be fed directly to the loudspeakers. As at present, the coded format (C-format) on disc or radio is different from the signal format used in the studio, and may be chosen to be compatible with most current 'matrix' or 'discrete' coding methods. The D-format signal in the listener's home will depend on the cost and complexity of his decoder but can be chosen for results greatly improving current methods.

To summarise, present studio quadraphonic processing equipment, designed for the so-called 'discrete' effect, gives very poor and unstable phantom sound images. A new generation of systems (called kernel systems) and associated studio signal processing equipment are under development in various countries to overcome these problems. A new approach based on the study[10] of the human hearing system has isolated those systems capable of results far more effective than discrete systems, which suggests which systems should be used and how they should be 'decoded' via four or more speakers for best possible results both horizontally and for sound in a three dimensional space including all directions and all distances. Studio signal handling formats designed to minimise problems and maximise inter-equipment compatibility are at the moment being studied and finalised, so as to avoid the mess that lack of standardisation would cause.

The technical approach outlined in the above will hopefully break the impasse caused by 'quadraphonic' approaches which could not even give reasonable results on master tape, barring 'interior' and the four corner positions, except if the listener were precisely central. The approach has been conceived so as to permit the creative producer a freedom of sound control not existing in 'discrete' approaches, including the full dimensions of three-dimensional space, 'interior' in-the-head effects, closeness, and the ability to move and modify whole fields of sound as well as individual sounds. These same features permit 'realistic' ambient concert hall recordings to be made with improved realism (reckoned by some to be as great a step beyond the best quadraphony as that is over the best stereo), and with a flexibility of control not hitherto possible.

While the range of possibilities are such that they are never likely to be exhausted, the technical understanding of psycho-acoustic kernel systems is now sufficient to point the way to removing many of the 'bugs' that have seemed inevitable and incurable in previous approaches.

References

1) K de Boer. 'A remarkable phenomenon with stereophonic sound reproduction' *Philips Technical Review*, vol 9, p8-13 (1947).

2) H A M Clark, G F Dutton, P B Vanderlyn. 'The "stereosonic" recording and reproducing system', *IRE Trans on Audio*, 1957, p6-111.

3) H D Harwood. 'Stereophonic Image Sharpness', *Wireless World*, p207-11, July 1968.

4) A D Blumlein. UK Patent 394325 (December 14, 1931).

5) M A Gerzon. 'The Principles of Quadraphonic Recording' (two parts), *Studio Sound*, Aug and Sept 1970.

6) O Kosaka, E Sato, T Nakayama. ‘Sound Image Localisation Muitichannel Matrix Reproduction’, *J Audio Eng Soc*, vol 20, p542-8 (Sept 1972).

7) P Scheiber. ‘Analysing Phase-Amplitude Matrices’, *J Audio Eng Soc*, vol 19, p835-9 (Nov 1971).

8) M A Gerzon. ‘Periphony: With-Height Sound Reproduction’, *J Audio Eng Soc*, vol 21, p2-10 (Jan/Feb 1973).

9) M A Gerzon. ‘Experimental Tetrahedral Recording’ (3 parts), *Studio Sound*, Aug, Sept, Oct 1971.

10) M A Gerzon. ‘Criteres Psychoacoustiques relatifs a la Conception des Systemes Matriciels et Discrets en Tetraphonie’ delivered at the Festival International du Son, Paris, 16th March, 1974, and published in its Journal *Conferences des Journees d’Etudes, Festival International du Son, 1974.*